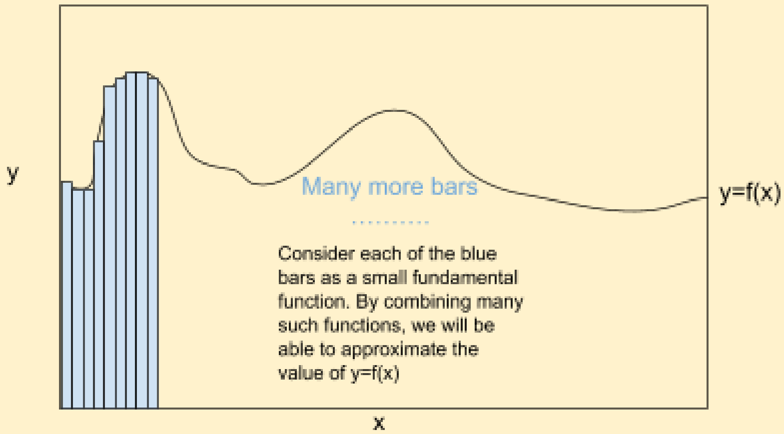
PadhAI: Representation Power of Functions

One Fourth Labs

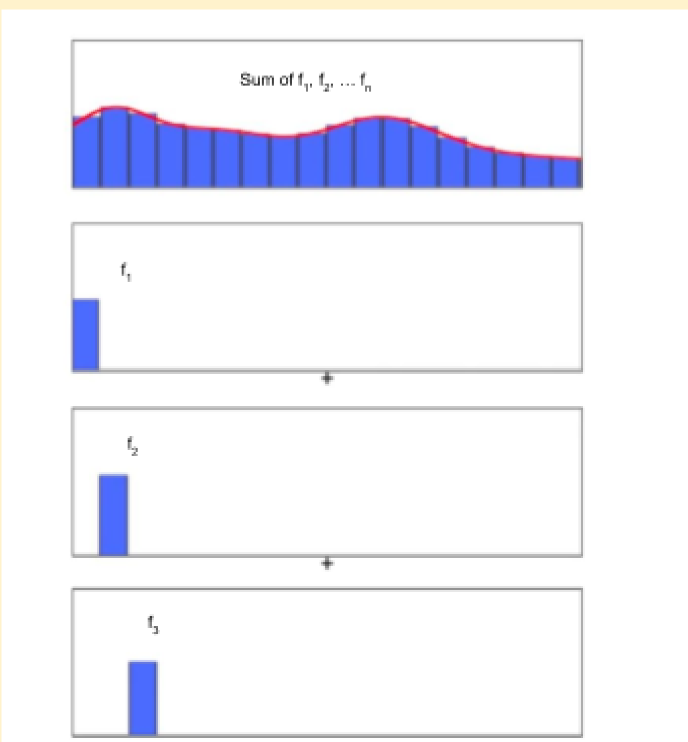
Illustrative proof of Universal Approximation Theorem

The representation power of deep neural networks

1. Consider the function $y = f(x)$, we want to obtain $\hat{f}(x)$ such that the two functions are are almost equal
2. However, creating a $\hat{f}(x)$ in one go is a daunting task
3. So, we can revisit our old analogy of building with bricks, where we represented a complex function as a combination of simple units
4. Consider the following illustration

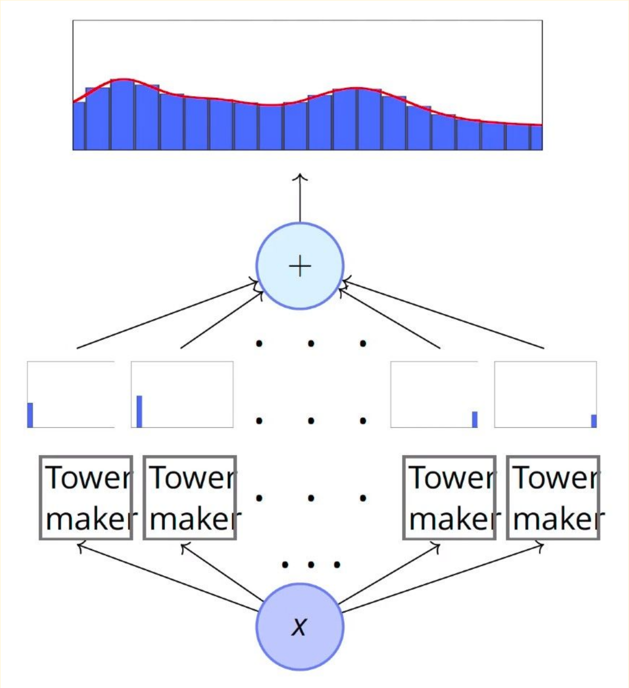


5. Here, the thinner the bar/tower, the better the approximation, because of less wasted space under/over the curve
6. Another illustration

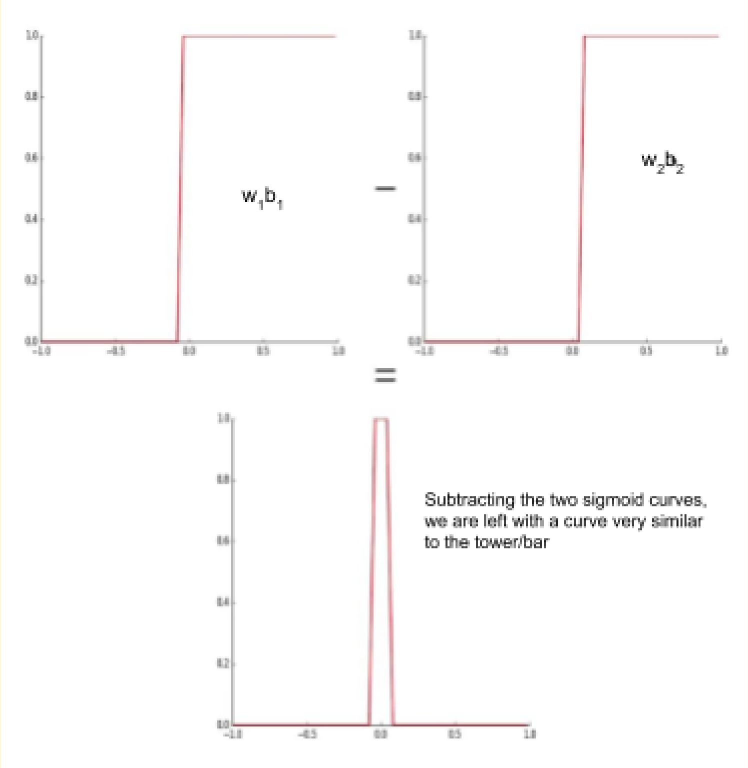


7. How does this tie back to the Sigmoid function

8. Consider the functions required to create these individual towers/bars



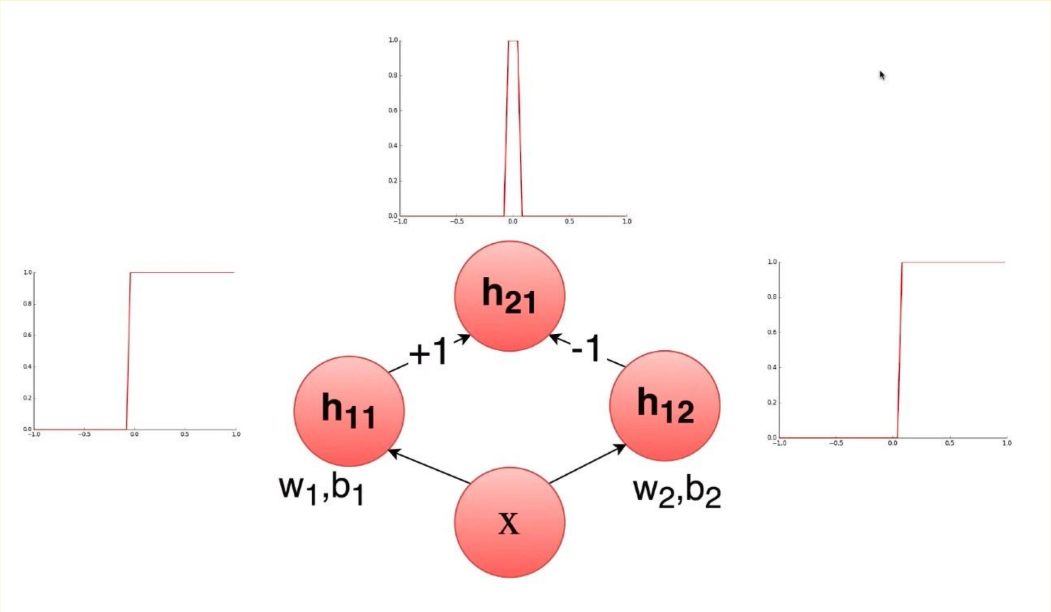
9. Let's see how the tower maker function is connected to the sigmoid function
10. In the sigmoid function, w is directly proportional to the sharpness of the curve and b shifts the horizontal position of the threshold. Consider subtraction between two sigmoid functions



PadhAI: Representation Power of Functions

One Fourth Labs

11. Neural network representation of sigmoid subtraction



12. With a network of many neurons, we will be able to create several towers/bars. These can then combine to approximate to any kind of function.